Moss diversity: A molecular phylogenetic analysis of genera

CYMON J. COX,1,3,5 BERNARD GOFFINET,2 NORMAN J. WICKETT2,4, SANDRA B. BOLES1 & A. JONATHAN SHAW1
1 Department of Biology, Duke University, Durham, North Carolina 27708, U.S.A.
2 Department of Ecology and Evolutionary Biology, University of Connecticut, Connecticut 06269, U.S.A.
3 Current address: Centro de Ciências do Mar (CCMAR), Universidade do Algarve, 8005-139 Faro, Portugal.
4Current address: Department of Biology, Pennsylvania State University, Pennsylvania 16802, U.S.A.
5 Author for correspondence (cymon.cox@googlemail.com)

Abstract

In this study we present phylogenetic and molecular phylogenetic diversity analyses of moss taxa from a total of 655 genera of mosses. Three loci were sampled: chloroplast ribosomal small protein 4, the intronic region of the mitochondrial NADH dehydrogenase subunit 5, and partial sequences of the nuclear 26S ribosomal RNA. Maximum likelihood and Bayesian phylogenetic analyses were performed on individual loci and on multilocus data sets. A measure of phylogenetic diversity was calculated and constrained among major lineages of mosses. We reveal many instances of incongruence among genomic partitions, but, overall, our analyses describe relationships largely congruent with previous studies of the major groups of mosses. Moreover, our greater sampling highlights the possible non-monophyly of many taxonomic families, particularly in the haplolepideous and pleurocarpous mosses. Comparisons of taxic and phylogenetic diversity among genera indicate that the Dicranidae (haplolepideous taxa) include about 15% of moss genera, but nearly 30% of the phylogenetic diversity. By contrast, the Hypnanae (hypnalian pleurocarps) contain about 45% of moss genera, but a lower percentage of phylogenetic diversity. Agreement between numbers of genera and phylogenetic diversity within other moss clades are remarkably consistent.

Key words: mosses, Bryopsida, phylogeny, phylogenetic diversity, Bayesian phylogenetic analysis

Introduction

Mosses (Bryophyta) are a diverse clade with over 12,700 species in more than 800 genera (Crosby et al. 1999) and are conspicuous floristic components in all terrestrial habitats suitable for plant growth — from the cold and barely hospitable antarctic to lush tropical rainforests. Mosses have many ecological functions, and one group, the peat mosses (*Sphagnum*), perform an important role in the global biogeochemical cycling of carbon where they represent a large reservoir of sequestered carbon in the form of northern boreal peatlands (Wieder & Vitt 2006).

Together with the other two bryophyte lineages, the liverworts (Marchantiophyta) and hornworts (Anthocerotophyta), and the vascular plants (Tracheophyta), mosses are one of the primary lineages of land plants (embryophytes) that arose *circa* 470 MYA, after the colonisation of land by an ancestor most closely related to modern-day charophycean algae (Lewis & McCourt 2004). The transition of plants from water to land was accompanied by major morphological, developmental, and physiological innovations most markedly expressed by the intercalation of a multicellular diploid phase into the life cycle (Kenrick 1994, Bateman et al. 1998, Graham et al. 2000, Renzaglia et al. 2000, Hemsley & Poole 2004). Based on extant representatives, the
charophycean green algal ancestor of all embryophyte land plants is assumed to have had a haplobiontic life cycle with a single, haploid, multicellular stage, whereas the bryophytes and tracheophytes exhibit a marked alternation of generations with a diplobiontic life-cycle. However, unlike tracheophytes, in which the dominant life cycle stage is the diploid sporophyte, the three bryophyte groups are characterized by dominant, perennial gametophytic stages, with relatively small, unbranched sporophytes that remain permanently attached to the maternal gametophytes.

The description and classification of moss diversity using microscopic characters was initiated by Johann Hedwig who recognised 35 moss genera based upon characters of the sporophyte peristome and gametophyte sexuality (Hedwig 1801). Thereafter, most classifications in the eighteenth century followed the system of Bridel-Brideri (1826–27) who emphasised the position of the perichaetium (the female inflorescence), and, consequently, the position of the sporophyte, as either terminal (acrocarpous) or lateral (pleurocarpous) on gametophyte shoot modules. Using this system, Bridel-Brideri described 91 acrocarpous and 31 pleurocarpous genera. Modern taxonomic treatments were initiated following the observations of peristome development and structure by Philibert (1884–1902). Classificatory systems emphasizing peristome architecture were established in the influential floras of Fleischer (1904–23) and Brotherus (1924, 1925) and remained dominant until the application of explicit evolutionary methodologies and the use of molecular phylogenies in the recent past (Buck & Goffinet 2000, Goffinet & Buck 2004, Goffinet et al. 2009). Today, approximately 845–866 genera of mosses are recognized (Crosby et al. 1999, Buck & Goffinet 2000).

In this study we describe the phylogeny and molecular diversity of mosses, focusing on genera as our operational taxonomic units. Although genera circumscribe an arbitrary degree of morphological diversity, they nevertheless provide a convenient boundary for the partitioning and comparison of morphological and molecular diversity. We have sampled 659 genera and analysed exemplar nucleotide sequences from the nuclear, plastid, and mitochondrial genomes.

### TABLE 1: Numbers of orders, families, genera, and species.

<table>
<thead>
<tr>
<th></th>
<th>Crosby et al. 1999</th>
<th>Buck &amp; Goffinet 2000</th>
<th>Goffinet &amp; Buck 2004</th>
</tr>
</thead>
<tbody>
<tr>
<td>orders</td>
<td>n/a</td>
<td>23</td>
<td>27</td>
</tr>
<tr>
<td>families</td>
<td>119</td>
<td>116</td>
<td>112</td>
</tr>
<tr>
<td>genera</td>
<td>845 (+56 synonyms)</td>
<td>866</td>
<td>866</td>
</tr>
<tr>
<td>species</td>
<td>12711</td>
<td>n/a</td>
<td>n/a</td>
</tr>
</tbody>
</table>

### Materials and Methods

**Taxon and locus sampling.** A total of 657 species representing 655 genera (two species of Takakia (T. lepidozoioides and T. ceratophylla) and two species of Leptotheca (L. boliviana and L. gaudichaudii) were sampled). Dried tissue was obtained mainly from the herbaria of the Missouri and New York Botanical Gardens (MO and NY, respectively). Wherever possible, material was selected from specimens determined by specialists in their respective taxonomic groups — these taxonomic assignments were not reassessed in light of the phylogenetic analyses conducted because of the large numbers of taxa sampled. Three loci were sequenced, one from each of the three genomic compartments; namely, the chloroplast small ribosomal protein 4 (rps4) gene, the intron and partial gene sequences of the mitochondrial NADH dehydrogenase subunit 5 (nad5), and partial sequences of the nuclear 26S ribosomal RNA (nuc26S). Sequence data were also obtained from NCBI GenBank when available for the same species. Taxon identities, nomenclatural authorship, voucher specimen information, and NCBI GenBank accession numbers are presented in Supplemental Information (S.I.) Table S.1. The precise root node of the mosses and hence the appropriate moss outgroup to the remaining mosses is somewhat uncertain. Newton et al. (2000) and Cox et al. (2004)
both resolved *Sphagnum* plus *Takakia* as the sister-group to other mosses, though without strong support from their preferred analyses. Here, the genus *Sphagnum* was used as an outgroup to polarize relationships (see Cox et al. 2004), with the understanding that *Sphagnum* plus *Takakia* may in fact form the lineage sister to all other mosses.

**DNA sequencing and alignment**

Total genomic DNA was extracted using the method of Edwards et al. (1991), or a standard CTAB procedure (Doyle & Doyle 1987), with subsequent cleaning using the Wizard DNA Clean-up Kit (Promega). Amplification and sequencing primers follow Cox et al. (2004) for the three DNA regions. For each taxon and sequenced DNA region, forward (5'-3') and reverse (3'-5') sequences were assembled and checked for inaccurate base calling using Sequencher (vers. 4.1, Gene Codes Corp.). Consensus sequences were aligned using ClustalW (Thompson et al. 1994) and manually adjusted using SeaView (vers. 3.2; Galtier, et al. 1996). Regions of ambiguous alignment and incomplete data (i.e., at the beginning and end of sequences) were identified and excluded from subsequent analyses.

**Phylogenetic analyses**

The best-fitting substitution model for the first, second, and third codon positions of the *rps*4 locus, and the *nad*5 and *nuc26S* regions, were assessed using Modeltest 3.06 (Posada & Crandall 1998) in conjunction with PAUP 4b10 (Swofford 2000). Two Bayesian Markov Chain Monte Carlo (MCMC) analyses were conducted using ‘mpi’ (parallel processing) version of MrBayes 3.0B4 (Ronquist & Huelsenbeck 2003) for each locus with parameter sampling every 500 generations for a total of 10,000,000 generations. The analyses were conducted with a single nucleotide substitution model for the *nuc26S* and *nad*5 loci, and a separate sub-model for each of the codon positions of the *rps*4 gene (i.e. a model data-heterogeneous analysis), with each model including the parameters as determined by the Modeltest analyses. Each MCMC analysis included four chains, one cold and three heated, and default prior distributions and proposal rates for each parameter value. The analyses were started from a user-defined starting tree (due to a limitation in the MrBayes 3.0B4 software which caused errors if random trees were used as starting trees) determined in PAUP (Swofford 2002) by the neighbor-joining algorithm with maximum likelihood distances under a GTR model. For each analysis the ‘burn-in’ period before the MCMC reached stationarity was determined by plotting the likelihood through time using the plotting program GNUPlot (Williams et al. 2004); stationarity was assumed to have been reached when the curve plateaued. Tree sets from the posterior distribution of the two independent runs for each locus were concatenated to form the sample of trees assumed to be randomly sampled from the posterior probability distribution.

Topological conflicts among loci were detected at a 95% posterior probability threshold, by comparison of the 50% majority-rule consensus trees of the trees recovered from the posterior probability distribution. The minimum number of conflicting taxa were pruned from the individual locus trees until no topological conflict remained between loci, with the locus comparison order of *rps*4 versus *nad*5, and *rps*4+*nad*5 versus *nuc26S*. After pruning conflicting taxa, matrices were combined to construct two concatenated matrices; namely, *rps*4/*nad*5 and *rps*4/*nad*5/*nuc26S*. The combined matrices, *rps*4/*nad*5 and *rps*4/*nad*5/*nuc26S, were pruned of all taxa that did not have all loci present. For both combined matrices, and for each locus and codon site partition in the *rps*4 gene, the appropriate model of nucleotide substitution was re-calculated using Modeltest. Data-heterogeneous MCMC analyses were performed on each combined matrix, with the substitution model of each character partition set to that recovered from the Modeltest analyses. Further details of the combined analyses are as described for the individual locus analyses.

Maximum likelihood (ML) analyses on individual gene partitions and combined data sets were performed with PhyML (Guindon & Gascuel 2003). Ten trees were randomly selected from the 95% confidence interval of the Bayesian MCMC analyses and the branch lengths optimised in PAUP under the model and parameter values found optimal for the partition by Modeltest. The 10 trees were then used as starting-trees and analysed by PhyML with the same optimal model but with parameter values estimated during the analyses. Because
PhyML only conducts homogeneous likelihood analyses, for the rps4 partition and the combined data sets, Modeltest was used to find the optimal model (and parameter values for the initial branch length calculations) under a single model prior to conducting the PhyML analyses.

Phylogenetic diversity
Phylogenetic diversity (PD - Faith 1992, 1994) is a measure proportional to the length of the minimum-spanning tree connecting a group of taxa. PD was calculated on the rps4/nad5/nuc26S data set for the following taxonomic groups (sensu Buck & Goffinet 2000): Polytrichiales, Funariales, Grimmiales, Pottiaceae, Bryaceae and Mniaceae, Bartramiaceae, Orthotrichales, Pteridiolemales, Dicranidae (Haplolepideae), and Hypnaceae (pleurocarpous mosses). PD values for each locus partition and for the combined data set for each taxonomic group were derived from the most likely tree obtained from the PhyML analyses and compared to the groups’ taxic diversity in terms of numbers of genera (Buck & Goffinet 2000).

Results
Single locus data sets
A total of 649, 634, and 578 sequences of the chloroplast rps4, mitochondrial nad5, and nuc26S loci, respectively, were newly generated during this project or were obtained from NCBI GenBank. Data characteristics of the sequences are given in Table 2. All data partitions, including individual loci, were found to be most appropriately described by a general time-reversible model of nucleotide substitution with gamma distributed rates of substitution among sites and a proportion of invariant characters (GTR+I+Γ: Rodriguez et al. 1990)(Supplementary Information Table S.2.1). MCMC analyses of the rps4, nad5, and nuc26S data sets each resulted in 30,415 unique topologies within the 95% posterior distributions after removal of 4,000 trees as the burn-in phase of each run and concatenation of the trees from both runs of each data set. The similarity in the number of unique topologies recovered by each analysis is because in each analysis no single topology was ever sampled twice, and the same numbers of topologies were discarded as burnin. The 50% majority-rule consensus trees of the combined tree set (assumed to be sampled from the posterior probability distribution) of each of the rps4, nad5, and nuc26S analyses is given in S.I. Figures S.2.1, S.2.2, and S.2.3, respectively. The optimal maximum likelihood tree found by PhyML of the rps4, nad5, and nuc26S data sets is presented in S.I. Figures S.2.4, S.2.5, and S.2.6, respectively.

Topological incongruence
A total of 284 topological conflicts were observed, at the threshold of 95% posterior probability support for node incongruence, between the rps4 and nad5 MCMC analyses. A total of 25 taxa were removed to eliminate all conflicts: these taxa are indicated in the S.I. (Materials and Methods). Of particular interest are a group of eight taxa, namely, Bruchia drummondii (Bruchiaceae), Calymperes lonchophyllum (Calymperaceae), Bellibarbula recurva (Pottiaceae), Campylopodium euphorocladium (Dicranaceae), Bryohumbertia filifolia (Dicranaceae), Bryomanginia saintpierrei (Ditrichaceae), Reimersia inconspicua (Pottiaceae), and Timmia megapolitana (Timmiaceae). In the rps4 phylogeny, the haplolepideous taxa (all the above taxa with the exception of Timnia) fall in well-supported groups close to members of their respective taxonomic families. Timnia is placed outside of the haplolepideae and close to the Encalyptaceae. However, in the nad5 tree all eight taxa form a monophyletic group with 100% posterior probability either to the exclusion of all other haplolepideous taxa (S.I. Figure S.2.2) or toward the base of haplolepideae (S.I. Figure S.2.5) but in neither case close to the Funariaceae. Moreover, the nucleotide sequences are remarkably similar with most sequences sharing similar polymorphic sites. These data characteristics suggest that these haplolepideous nad5 sequences are the result of human error in collecting the data: however, there are unique differences among the sequences and this has not been tested with additional sequencing.
Between the rps4/nad5 and nuc26S MCMC analyses, a total of 102 node conflicts were observed, at the threshold of 95% posterior probability support for topological incongruence. A total of 62 taxa were removed to eliminate conflicts among the two partitions (S.I.: Materials and Methods). Of particular note was the conflict between the positions of the Ptychomniales in the topologies. In the rps4/nad5 tree the Ptychomniales were the sister-group to all other pleurocarpous mosses (Buck et al. 2004), whereas in the nuc26S tree they were allied with Aulacomniaceae, Phyllodrepaniaceae, Bryaceae, Mniaceae, and a few members of the Rhizogoniaceae (all non-pleurocarpous families). The Ptychomniales were removed from the combined rps4/nad5/nuc26S data set. In addition, the acrocarpous order Splachnales was removed as the order was resolved as a monophyletic unit among the pleurocarpous mosses in the nuc26S data set in contrast to the traditional placement among the diplolepideous-alternate peristomate mosses leading in a grade to the pleurocarpous mosses, as was resolved by the rps4/nad5 data set.

**Multilocus data sets**

The combined rps4/nad5 data set consisted of 603 taxa after removal of 25 taxa in conflict between the partitions and a further 31 taxa for which only one of the loci was available. The aligned matrix included 3718 sites of which 2105 sites were excluded due to ambiguous alignment or missing data at the ends of the sequences. The remaining 1613 sites included 1002 parsimony informative sites. The combined rps4/nad5/nuc26s data set consisted of 477 taxa after removal of 62 taxa in conflict between the data sets and 103 taxa that had missing data. Both combined data sets were most-appropriately described by a GTR+I+Γ substitution model (S.I. Table S.2.1). Heterogeneous MCMC analysis of the rps4/nad5 data set resulted in 30,145 unique topologies within the 95% confidence interval of the posterior probability distribution, while similar analyses for the rps4/nad5/nuc26s data set resulted in 22,797 unique trees. The 50% majority-rule consensus trees of the topologies from the posterior probability distribution are presented in S.I. Figure S.2.7 (rps4/nad5) and Figure S.2.8 (rps4/nad5/nuc26s). The optimal maximum likelihood tree found by PhyML of the rps4/nad5 and rps4/nad5/nuc26S data sets is presented in S.I. Figures S.2.9 & S.2.10, respectively.

**Phylogenetic relationships**

Overall, relationships indicated by the three individual loci are remarkably similar, with the nematodontous mosses forming a grade at the base of the trees leading to a monophyletic arthrodontous moss clade. Within the arthrodontous mosses, the Encalyptales and Funariales tend to be sister to a haplolepideous/diplolepideous-alternate peristomate moss split (rps4) or of unresolved affinities (nad5, nuc26S). The haplolepideae are monophyletic (nad5 – minus the “Timmia” clade, see above) or lacking sufficient support to
contradict their monophyly (rps4, nuc26S). The monophyly of the diplolepideous-alternate moss clade is strongly supported by all loci, with, typically, a grade of acrocarpous families and orders leading to a monophyletic pleurocarpous mosses. With the exception of the nuc26S locus, the monophyly of the pleurocarpous mosses is strongly supported.

Although the trees presented here are informally rooted (outgroup rooting) with *Sphagnum*, we note that the rps4 and nad5 loci do not contradict previous studies that placed both *Sphagnum* and *Takakia* as the sister group to the remaining mosses (Newton et al. 2000, Cox et al. 2004), while nuc26S does not provide strong conflicting evidence. The placement of Andreaeales as a clade sister to (nad5), or a grade leading to (rps4), other mosses is again in agreement with these previous studies. The relationships concerning the Polytrichales, Buxbaumiales, Oedipodiales, Tetraphidales, and arthrodontous mosses are problematic. The rps4 data set has strong support (100% posterior probability (p.p.)) for the sister relationship of the Polytrichales, Buxbaumiales, Oedipodiales, Tetraphidales, and arthrodontous mosses are problematic. The rps4 data set has strong support (100% posterior probability (p.p.)) for the sister relationship of the Polytrichales, Buxbaumiales, Oedipodiales, Tetraphidales, and arthrodontous mosses are problematic. The rps4 data set has strong support (100% posterior probability (p.p.)) for the sister relationship of the Polytrichales, Buxbaumiales, Oedipodiales, Tetraphidales, and arthrodontous mosses are problematic. The rps4 data set has strong support (100% posterior probability (p.p.)) for the sister relationship of the Polytrichales, Buxbaumiales, Oedipodiales, Tetraphidales, and arthrodontous mosses are problematic. The rps4 data set has strong support (100% posterior probability (p.p.)) for the sister relationship of the Polytrichales, Buxbaumiales, Oedipodiales, Tetraphidales, and arthrodontous mosses are problematic.

Placement of the order Buxbauma as the sister-group to the arthrodontous mosses is resolved only in the nad5 and combined analyses; although the relationship does not receive significant support from the analyses. The sister-group relationship of Diphysciales to other arthrodontous mosses is well supported by both the rps4 and nad5 loci, and the combined analyses.

The relationship of the Funariales to the Encalyptales and Timmiales, to the exclusion of the Gigaspermaeae, is supported by the rps4 locus alone, with strong contradictory evidence provided by the nuc26S locus, which places Timmia in a clade with the haplolepideous taxa, Timmiella (Pottiaceae) and Distichium (Distichicaceae). The combined rps4/nad5 (i.e. without Timmia) data set strongly supports the clade uniting Funariales and Encalyptales (100% p.p.), with Gigaspermaeae as the immediate sister-group (75% p.p.). However, the support for these relationships is much reduced by the addition of the nuc26S data; 82% p.p. and 60% p.p. respectively. In the rps4/nad5 combined analyses the Funariales/Encalyptales clade is placed as the sister group to the haplolepideae, whereas in the rps4/nad5/nuc26S analyses the clade is resolved as sister to the haplolepideae plus Bryineae, however, in neither case is the relationship significantly supported.

The monophyly of the haplolepideous mosses is only supported by nad5 among the single-locus analysis (excepting the “Timmia” clade), but is maximally supported in combined analyses. In general, the relationships among haplolepideous taxa are well resolved and supported, although few traditional groupings are evident from the analyses. The enigmatic taxon, *Catosciopium nigritum*, classified in the Splachnaceae by Buck & Goffinet (2000), is resolved as the first-branching taxon in the haplolepideae clade of the rps4/nad5 tree, with a clade consisting of *Chrysoblastella* (Distichicaceae), *Luisierella* (Pottiaceae), *Distichium* (Distichicaceae), and *Timmia* (Pottiaceae) also forming an early branching lineage separated from the remaining haplolepideae. A clade consisting of *Scouleria* (Scouleriaceae), *Dieranowiesia* (Rhabdoweisiae), and *Drummondia* (Drummondiaceae) form the next branching lineage, followed by separation of *Bryoxiphium* (Bryoxiphiaeae) in the rps4/nad5 tree. However, in the three gene combined analysis *Bryoxiphium* forms a clade with *Drummondia* and *Scouleria* although the relationship is not well-supported. A well-supported clade is formed between members of the Grimmiaeae and Ptychomitriaceae (100% p.p.), although the latter appears to be paraphyletic with the grimmiaeae genera *Jaffueliobryum* and *Indusiella* derived from the Ptychomitriaceae. Sister to this clade are *Saelania* (Distichicaceae) and *Blindia*
(Seligeriaceae) respectively. The gymnostomous (lacking a peristome) taxon, *Pseudobraunia*, typically classified with the Hedwigiales, is also recognised as a member of the Grimmiales by each of the loci, with strong support (99% p.p.) for its sister relationship to *Dryopteris* in the *rps4/nad5* analyses.

The Grimmiales/Ptychomitriaceae clade is resolved (75% p.p.) as sister to a clade consisting of taxa from the Dicranaceae, Leucobryaceae, Ephemereaceae (*Micromitrium* only), and Archidiaceae. The remaining haplolepideae form a well-supported (100% p.p.) clade in the *rps4/nad5* data set, which is considerably weakened (89% p.p.) by the addition of the nuc26S data. Within this latter group, *Amphidium* (Rhabdoweisiaceae) forms the sister taxon to the remaining taxa, which themselves are divided into two well-supported lineages. Both of these latter clades are very heterogeneous in taxon composition, the first consisting of members of the Rhabdoweisiaceae, Fissidentaceae, Erpodaceae, Bruchiaceae, Ditrichaceae, Pottiaceae, Serportortellaceae, Calymperaceae, and Dicranaceae. The second clade consists of the majority of members of the Pottiaceae (100% p.p.), plus the remainder of the Dicranaceae, Dichmemonaceae, Rhabdoweisiaceae, Ditrichaceae, Rhachithecaceae, Erpodaceae, Bruchiaceae, and the monogeneric Wardiaceae.

The diplolepide-alternate peristomate mosses (Bryidae) are maximally supported as a monophyletic group in all individual and combined analyses. The Bartramiaeae, Hedwigiales, Mniaceae, Bryaceae, and Splachnales, all form well-supported groups in the *rps4/nad5* tree. In the *rps4/nad5/nuc26S* analyses, the Hedwigiales or Bartramiaeae are identified as the two first-branching lineages, with neither group supported as the earliest branching lineage. However, the removal of the Splachnales from the three gene combined analyses was necessary due to its positioning within the pleurocarpous taxa in the tree inferred from nuc26S, nevertheless, the *rps4* and *nad5* loci suggest the group may also be one of the earliest diverging lineages of the Bryidae. The Bryaceae plus Leptostomataceae and the Phyllodrepaniaceae plus Mniaceae both form a maximally supported monophyletic group, and together these four families form a clade (100% p.p.) sister to the remaining Bryidae. Of note is the inclusion of the genus *Pulchrinodus* in this clade, with good support (97% p.p.) for its sister relationship to the Leptostomataceae plus Bryaceae in the *rps4/nad5* trees (c.f. Stech *et al.* 2003).

The Orthotrichales, which themselves are a well-supported lineage (100% p.p.), are supported (97% p.p.) as the sister lineage to the Rhizogoniales and pleurocarpous mosses by the *rps4/nad5* analyses. The support is, however, significantly lessened by the addition of nuc26S data (80% p.p.).

The Rhizogoniales, as traditionally circumscribed, form a grade of taxa leading to the pleurocarpous mosses (Hypnaceae); three well-supported clades of rhizogoniaceeous taxa are resolved with the individual *rps4* and *nad5*, and the combined *rps4/nad5* analyses. The first indicates that the Orthodontiaceae (Bryales) are more closely related to *Leptothesca* and *Hymenodon* (Rhizogoniales) than to other member of the Bryales. The second clade includes *Rhizogonium, Goniobryum, Pyrrhobryum, Cyrtopodium*, all classified in the Rhizogoniales, plus *Calomnion* from the Calomniaceae – *Calomnion* and *Cyrtopodium* are maximally supported as sister groups. The third clade includes *Hypnodendron* (Hypnodendraceae), *Spiridens* (Spiridentaceae), *Cytopus* (Cytopodaceae), *Cyrtopodendron* (Cytopodaceae), *Pterobryella* (Pterobryellaceae), *Racopilum* (Racopilaceae), and *Braithwaitaea* (Trachylomataceae). Furthermore, the Bryalean family Aulacomniaceae is also part of this rhizogoniaceeous grade leading to the pleurocarps. Only the *nad5* locus provides strong evidence for the resolution order of these clades, while the *rps4* locus does not conflict with this set of relationships, and the nuc26S provides strikingly contradictory evidence, although of dubious validity (e.g. part of the Hypnodendraceae are nested within the Orthotrichales). In the *nad5* tree, the Orthodontiaceae are the first branching lineage followed by the branching of the Rhizogoniales, then the Aulacomniaceae, and finally the Hypnodendraceae which is the sister-group to the pleurocarpous mosses (taxonomy *sensu* Bell *et al.* 2007). In the *nad5* analyses, only the position of the Aulacomniaceae is uncertain, yet all relationships are significantly weakened by the addition of nuc26S data, excepting the placement of the Hypnodendraceae.

In the individual *rps4* and *nad5* analyses and the *rps4/nad5* combined analysis, the Ptychommiales (*sensu* Buck *et al.* 2004) are significantly supported (96% p.p. and 100% p.p., respectively) as the sister-group to all
other pleurocarpous mosses (c.f. Cox et al. 2004). In contrast, the nuc26S data place the order with the Aulacomniaceae and Phyllodrepaniaceae, and closely related to the Mniaceae and Bryaceae.

In none of the rps4, nad5, or combined rps4/nad5 analyses are the Hookeriales resolved as a monophyletic group, but rather they form two strongly supported and distinct groups. One consists of the two monogeneric families Climaciaceae (Climacium Regmatodontaceae (Regmatodon members of the Leskeaceae (Hypopterygiaceae, plus maximum support for Hypnalean monophyly. However, addition of the nuc26S data resolves a paraphyletic and (Plagiothecium supported clade with the monogeneric family Plagiotheciaceae (Mahua). The two monogeneric families Rhytidiaceae (Campyliaceae (Cyathophorella, and (Palustriella), and Eriodon (Brachytheciaceae). Many other examples could be given of similar conflict with traditional taxonomy.

The Hypnalean pleurocarpous mosses are unresolved as a monophyletic group by the rps4 locus, resolved as a clade by the nad5 locus (86% p.p.), and supported as polyphyletic by the nuc26S locus (but the Splachnales are strongly supported as derived from the hypnalean clade). The combined rps4/nad5 analyses provide maximum support for Hypnalean monophyly. However, addition of the nuc26S data resolves a paraphyletic (with Hypopterygiaceae plus Cyathophorella nested within) Hypnaleas, but without significant support.

Overall, support for lineages in the Hypnaleas was poor, but the few sizable clades that were resolved indicate that all the large Hypnalean families are non-monophyletic. In the rps4/nad5 combined analyses, a large brachytheciaceous clade is resolved (Bryoandersonia – Schwetschkea in S.I. Figure S.2.7) but also contains representatives from Hyocomiaceae (Neodolichomitrea), Stereophyllaceae (Entodontopsis), Campyiaceae (Hygrohypnum), Leskeaceae (Okamuraea, Schwetschkea), and Myriniaceae (Helicodontium). Similarly, the clade Saniona – Pseudocalliergen (99% p.p.) consists of the majority of Campyiaceae, but also includes members of the Hypnaleas (Campylophyllum, Taxiphyllopsis), Amblystegiaceae (Amblystegium, Hygroamblystegium, Orthotheciella, Vittia, Gradsteinia), Cratoneuronaceae (Cratoneuron), Helodiaceae (Palustriella), and Eriodon (Brachytheciaceae). Many other examples could be given of similar conflict with traditional taxonomy.

Some small families of the Hypnaleas are placed with support on the rps4/nad5 tree. The small family Rutenbergiaceae (Neorutenbergia, Rutenbergia, and Pseudocryphaea) is monophyletic (98% p.p.) and forms a clade (95% p.p.) with Trachyloma, the sole genus in the family Trachylomataceae (excepting Braithwaitaea), and Rhizofabronia (Fabroniaceae). The two monogeneric families Rhytidiaeae (Rhytidium) and Regmatodontaceae (Regmatodon) are well-supported (99% p.p.) as belonging to a clade that includes members of the Leskeaceae (Linbergia, Pseudoleskeella, Leskeadelphus, Pseudoleskeopsis) and Fabroniaceae (Dimerodontium). The monogeneric Rigidiaceae (Rigodium) are placed in a well-supported clade (99% p.p.) with Camptochaete, Lembophyllum, and Weymouthia, all members of the Lembophyllaceae. Isopterygiopsis and Mahua, both members of the Hypnaleas, along with Struckia from the Amblystegiaceae, form a maximally supported clade with the monogeneric family Plagiotheciaceae (Plagiothecium). The Fontinalaceae (Fontinalis, Brachylena, and Dichelyma) are monophyletic (100% p.p.) and forms a sister-group to Habrodon (Pterigynandraceae). The two monogeneric families Climaciaceae (Climacium) and Pleuroziopsiaceae (Pleuroziopsis) form a monophyletic group (100% p.p.). The Symphyodontaceae (Symphyodon, Dimorphochlodon, Chaetomitriopsis and Chaetomitrium) form a clade with Phyllophon and Bryocrumia, both members of the Hypnaleas. Finally, the shared ancestry between the monogeneric Theliaceae (Thelia) and Leptopterigynandrum (Leskeaceae), Schwetschkeopsis (Anomodontaceae), and Taxiphyllum (Hypnaleas) is maximally supported.

**Taxic and phylogenetic diversity**

Numbers of genera (taxic diversity) and phylogenetic diversity (PD) are compared for selected groups within the mosses in Table 3. We limit these comparisons to monophyletic groups. Phylogenetic diversity estimated from each of the three loci, and taxic diversity, are all highly correlated; correlation coefficients among all pairwise combinations of estimates are greater than 0.90 (statistics not shown). Thus, diversity estimates based on at least broad groupings recognized in the widely accepted classification scheme captures phylogenetic structure in the mosses.
Within that broad agreement, some consistent patterns of relationship between phylogenetic and taxic diversity estimates emerge. The Dicranidae (haplolepideous taxa), for example, include about 15% of moss genera, but nearly 30% of the phylogenetic diversity. The hypnalian pleurocarps (Hypnanae in Table 3), in contrast, contain about 45% of moss genera, but a lower percentage of phylogenetic diversity. One way to interpret this pattern is that genera within the Dicranidae are, on average, separated by a larger phylogenetic distance than are hypnalean genera. Agreement between numbers of genera and phylogenetic diversity within other moss clades are remarkably consistent.

**TABLE 3:** Phylogenetic and taxic diversity based on the best ML tree obtained from the PhyML analysis of the *rps4/nad5/nuc26s* combined data set.

<table>
<thead>
<tr>
<th>Taxon1</th>
<th>Genera2</th>
<th><em>rps4</em> (10.247)3</th>
<th><em>nad5</em> (4.059)</th>
<th><em>nuc26s</em> (5.163)</th>
<th><em>rps4/nad5/nuc26S</em> (7.539)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Polytrichales (11: 1.6%)</td>
<td>23</td>
<td>3.1% (0.320)</td>
<td>3.4% (0.136)</td>
<td>4.0% (0.206)</td>
<td>3.3% (0.250)</td>
</tr>
<tr>
<td>Funariales' (9: 1.4%)</td>
<td>24</td>
<td>3.3% (0.339)</td>
<td>3.5% (0.143)</td>
<td>6.4% (0.330)</td>
<td>4.1% (0.305)</td>
</tr>
<tr>
<td>Grimmiaaceae (7: 1.1%)</td>
<td>10</td>
<td>1.5% (0.157)</td>
<td>1.2% (0.048)</td>
<td>2.0% (0.104)</td>
<td>1.5% (0.113)</td>
</tr>
<tr>
<td>Pottiaceae (34: 5.2%)</td>
<td>82</td>
<td>7.2% (0.741)</td>
<td>8.2% (0.332)</td>
<td>5.3% (0.275)</td>
<td>7.1% (0.536)</td>
</tr>
<tr>
<td>“Bryum/Mnium”” (21: 3.2%)</td>
<td>28</td>
<td>5.2% (0.530)</td>
<td>2.9% (0.116)</td>
<td>6.1% (0.313)</td>
<td>4.6% (0.347)</td>
</tr>
<tr>
<td>Bartramiaceae (8: 1.2%)</td>
<td>10</td>
<td>1.3% (0.132)</td>
<td>2.8% (0.112)</td>
<td>2.7% (0.138)</td>
<td>2.2% (0.162)</td>
</tr>
<tr>
<td>Orthotrichales (13: 2.0%)</td>
<td>22</td>
<td>4.3% (0.443)</td>
<td>3.4% (0.137)</td>
<td>5.8% (0.299)</td>
<td>4.1% (0.314)</td>
</tr>
<tr>
<td>Dicranidae (101: 15.3%)</td>
<td>234</td>
<td>28.7% (2.941)</td>
<td>28.0% (1.137)</td>
<td>27.1% (1.400)</td>
<td>28.1% (2.120)</td>
</tr>
<tr>
<td>Bryidae (348: 52.9%)</td>
<td>91</td>
<td>55.2% (5.653)</td>
<td>47.3% (1.919)</td>
<td>56.9% (2.937)</td>
<td>53.1% (4.006)</td>
</tr>
<tr>
<td>Grimmiaceae (7: 1.1%)</td>
<td>10</td>
<td>1.5% (0.157)</td>
<td>1.2% (0.048)</td>
<td>2.0% (0.104)</td>
<td>1.5% (0.113)</td>
</tr>
<tr>
<td>Pottiaceae (34: 5.2%)</td>
<td>82</td>
<td>7.2% (0.741)</td>
<td>8.2% (0.332)</td>
<td>5.3% (0.275)</td>
<td>7.1% (0.536)</td>
</tr>
<tr>
<td>“Bryum/Mnium”” (21: 3.2%)</td>
<td>28</td>
<td>5.2% (0.530)</td>
<td>2.9% (0.116)</td>
<td>6.1% (0.313)</td>
<td>4.6% (0.347)</td>
</tr>
<tr>
<td>Bartramiaceae (8: 1.2%)</td>
<td>10</td>
<td>1.3% (0.132)</td>
<td>2.8% (0.112)</td>
<td>2.7% (0.138)</td>
<td>2.2% (0.162)</td>
</tr>
<tr>
<td>Orthotrichales (13: 2.0%)</td>
<td>22</td>
<td>4.3% (0.443)</td>
<td>3.4% (0.137)</td>
<td>5.8% (0.299)</td>
<td>4.1% (0.314)</td>
</tr>
<tr>
<td>Dicranidae (101: 15.3%)</td>
<td>234</td>
<td>28.7% (2.941)</td>
<td>28.0% (1.137)</td>
<td>27.1% (1.400)</td>
<td>28.1% (2.120)</td>
</tr>
<tr>
<td>Bryidae (348: 52.9%)</td>
<td>91</td>
<td>55.2% (5.653)</td>
<td>47.3% (1.919)</td>
<td>56.9% (2.937)</td>
<td>53.1% (4.006)</td>
</tr>
<tr>
<td>Grimmiaceae (7: 1.1%)</td>
<td>10</td>
<td>1.5% (0.157)</td>
<td>1.2% (0.048)</td>
<td>2.0% (0.104)</td>
<td>1.5% (0.113)</td>
</tr>
<tr>
<td>Pottiaceae (34: 5.2%)</td>
<td>82</td>
<td>7.2% (0.741)</td>
<td>8.2% (0.332)</td>
<td>5.3% (0.275)</td>
<td>7.1% (0.536)</td>
</tr>
<tr>
<td>“Bryum/Mnium”” (21: 3.2%)</td>
<td>28</td>
<td>5.2% (0.530)</td>
<td>2.9% (0.116)</td>
<td>6.1% (0.313)</td>
<td>4.6% (0.347)</td>
</tr>
<tr>
<td>Bartramiaceae (8: 1.2%)</td>
<td>10</td>
<td>1.3% (0.132)</td>
<td>2.8% (0.112)</td>
<td>2.7% (0.138)</td>
<td>2.2% (0.162)</td>
</tr>
<tr>
<td>Orthotrichales (13: 2.0%)</td>
<td>22</td>
<td>4.3% (0.443)</td>
<td>3.4% (0.137)</td>
<td>5.8% (0.299)</td>
<td>4.1% (0.314)</td>
</tr>
<tr>
<td>Dicranidae (101: 15.3%)</td>
<td>234</td>
<td>28.7% (2.941)</td>
<td>28.0% (1.137)</td>
<td>27.1% (1.400)</td>
<td>28.1% (2.120)</td>
</tr>
<tr>
<td>Bryidae (348: 52.9%)</td>
<td>91</td>
<td>55.2% (5.653)</td>
<td>47.3% (1.919)</td>
<td>56.9% (2.937)</td>
<td>53.1% (4.006)</td>
</tr>
<tr>
<td>Grimmiaceae (7: 1.1%)</td>
<td>10</td>
<td>1.5% (0.157)</td>
<td>1.2% (0.048)</td>
<td>2.0% (0.104)</td>
<td>1.5% (0.113)</td>
</tr>
<tr>
<td>Pottiaceae (34: 5.2%)</td>
<td>82</td>
<td>7.2% (0.741)</td>
<td>8.2% (0.332)</td>
<td>5.3% (0.275)</td>
<td>7.1% (0.536)</td>
</tr>
<tr>
<td>“Bryum/Mnium”” (21: 3.2%)</td>
<td>28</td>
<td>5.2% (0.530)</td>
<td>2.9% (0.116)</td>
<td>6.1% (0.313)</td>
<td>4.6% (0.347)</td>
</tr>
</tbody>
</table>

1. Taxonomic group and number of taxa (with percentages) included in the analysis (*rps4/nad5/nuc26s* data set).
3. Total tree length is indicated with the locus in each column.
5. Includes the Bryaceae, Mniaceae, and Leptostomataceae.

**Discussion**

The broad pattern of phylogenetic relationships among the mosses we present here is in general agreement with our current understanding of evolutionary relationships based on molecular data (Newton *et al.* 2000, Cox *et al.* 2004, Bell *et al.* 2007, Wahrmund *et al.* 2010). Our combined analysis (Figure 1A) depicts the Andreaeopsida as the sister-group of the peristome mosses (plus *Oedipodium*, which is gymnostomatous) with the nematodontous mosses forming a grade of taxa leading to the arthrodontous mosses. Within the latter, the haplolepideous are derived from mosses with a diplolepideous peristome and pleurocarpous mosses are derived from acrocarpous diplolepideous mosses. Among the primary phylogenetic divisions, no divergences gain support here that have not previously been firmly established, and some are without support where previously support was evident (cf. Cox *et al.* 2004). These differences in levels of support may reflect the relative paucity of data per taxon in our current analyses where some earlier analyses have focused on obtaining greater amounts of data per taxon at the expense of taxon sampling density. On the other hand, the increased taxon sampling here may have reduced support where in previous analyses it was erroneously present due to sparse taxon sampling. The relative importance of dense taxon sampling versus obtaining larger amounts of data per taxon is a long debated topic in phylogenetics (e.g. Graybeal 1998, Rosenberg & Kumar 2001, Zwickl & Hillis 2002), but is often discussed with regard to phylogenetic analyses as a whole. However,
FIGURE 1. The optimal maximum likelihood tree of the combined rps4/nadS/nuc26S data set under the general time-reversible model with site rate variation (GTR+I+Γ: -ln likelihood = 54575.7; optimal parameters: A/C = 1.1166, A/G = 5.0600, A/T = 0.2165, C/G = 1.0611, C/T = 5.6457, G/T = 1(fixed); f(A) = 0.3360, f(C) = 0.1836, f(G) = 0.1781, f(T) = 0.3027; α = 0.5730; pinvar = 0.3410). Nodes support by >95% posterior probability are highlighted. Taxonomic labels are applied to clades at the lowest rank applicable to the clade following the classification of Buck and Goffinet (2000): A) acrocarpous mosses - the two exemplars marked with an asterisk were originally labeled *Hymenostylis recurvirostre* and *Pottia truncata* and subsequently identified as *Ceratodon purpureus*, B) pleurocarpous mosses - part 1, C) pleurocarpous mosses - part 2.
taxon-dense and data-poor trees are not necessarily more accurate than taxon-sparse and data-rich trees per se, and visa versa, for a given set of taxa. The issue is perhaps most beneficially viewed with regard to specific nodes in the tree, which are most likely to require a balance of consideration for both these factors (Hillis et al. 2003). That is, while some nodes are relatively easy to resolve with either taxon-sparse or data-poor analyses, difficult to resolve nodes may require larger taxon samples, perhaps to break-up long-branches in the vicinity, or larger amounts of data to resolve short-branches in the tree, or, most probably, a balance of both of these factors. The analyses we present here are very taxon-dense in particular parts of the tree (e.g. the pleurocarpous mosses) while being relatively taxon-poor in others (e.g. within the nematodontous grade). Our aim was to provide an overview of molecular diversity among the mosses from each of the three genomic compartments, and as such we emphasized taxon sampling at the expense of data sampling (though clearly we consider the amount of data we collected to be enough to make sufficiently accurate trees), but we do not consider these trees to be more accurate with regard to any particular divergence merely due to the increased taxon sampling. It is, however, noteworthy that the analyses presented here are for the most part congruent with gene-rich studies regarding many of the deep-level relationships we identify.

The many incongruencies between genes in our analyses and, more generally, between our study and previous studies, highlight the necessity of identifying the sources of conflicting signals in phylogenetic analyses. The incongruencies may be 'real' in the sense that they reflect different evolutionary histories between the nuclear and two cytoplasmic genomes, via mechanisms such as lateral gene transfer (Bergthorsson et al. 2004) or differential retention in lineages of copies of duplicated genes (where paralogs can wrongly be assumed to be orthologs), but such conclusions are often the explanation of last resort (because they have such profound biological implications) when all other possibilities have been eliminated. Alternatively, as is most likely here, they are the result of analytical artifacts introduced by differing taxon samples and applied methodologies, although human error (data collection and manipulation errors) is often difficult to discount. Of particular concern here is the failure of the applied models to adequately reflect the underlying processes of molecular evolution that gave rise to the taxon gene sequences. Substitution models are typically chosen as best-fitting from a selected few possible models (e.g. using Modeltest), but rarely are they tested for adequacy — a model may be the best-fitting of those available but at the same time fail to reflect the underlying substitution processes adequately (Bollback 2001). Furthermore, in a standard model-based methodology substitution models are applied such that the substitution process described by the instantaneous rate matrix (Q; Swofford et al. 1996) remains constant across the entire tree within a particular data partition (be it gene, codon-site, or how ever the data are divided). Nevertheless, as we go deeper in the tree below species-level diversity it is increasingly unlikely that constancy of this substitution process holds, regardless of changes in overall rate of change (i.e. branch length variation). The older a divergence the longer have its descendant lineages had time to diverge in substitution dynamics — often, we need only look at taxon base compositions to confirm that indeed the substitution process itself has changed through time. Failing to model these changes can have detrimental effects on phylogenetic accuracy (Foster 2004), but the extent to which particular analyses are affected is usually unknown. In our analyses, we constructed a data set with the largest possible taxon sample containing all three genes, where those genes exhibited no evidence of severe incongruence (as evidenced by < 95% p.p. of conflicting resolution of lineages). By doing so we attempted to at least minimise the effects associated with intra-taxon conflict, whatever its origin, between data partitions with the expectation that inter-taxon resolution and statistical support would be maximised in the tree.

**Early-branching peristomate lineages**

The relationships among the early-diverging peristomate moss lineages have proven especially difficult to resolve with confidence (Newton et al. 2000, Hyvönen et al. 2004, Cox et al. 2004, Wahrmund et al. 2010), and our analyses here provide only a little further insight. The two nematodontous lineages, Tetraphidales and Polytrichales, compose a monophyletic group only when inferences were made based on the mitochondrial genes *nad2* and *nad5* (78% maximum likelihood bootstrap; Beckert et al. 2001). In the current analysis of the *nad5* gene (S.I. Figure S.2.2) we found no evidence of a sister-group relationship between the two groups (the
nad2 gene in the Beckert et al. study did not provide any statistical support for the monophyly of the nematodontous mosses). The rps4 gene alone strongly supported the Tetraphidales as the earliest diverging lineage of peristomate mosses, with the gymnostomous taxon Oedipodium and the (at least partially) arthrodontous taxon Buxbaumia forming a clade with the Polytrichales (S.I. Figure S.2.1). By contrast, previous analyses of the rps4 gene have placed Oedipodium as the earliest diverging lineage, with the Tetraphidales and Buxbaumia forming a clade sister to the Polytrichales, although these relationships received no statistical support (Goffinet et al. 2001). When combined with the nad5 and nuc26S data, the statistical support from our rps4 analyses for the resolution of the Tetraphidales as the earliest-diverging lineage was eroded (i.e. <95% pp. Figure 1A), as was the support for the placement of Oedipodium and Buxbaumia within a clade with the Polytrichales. Inferences from additional loci combined with morphological data also failed to unambiguously resolve the affinities of the Tetraphidales (Hyvönen et al. 2004). Given the ambiguity, the origin of the arthrodontous peristome, its derivation from a nematodontous one, and even the fundamental homology between the peristome of the Tetraphidales and Polytrichales, remains obscure.

Early-branching arthrodontous lineages

The Diphysciales are well established as the first diverging lineage of arthrodontous mosses (Newton et al. 2000, Beckert et al. 2001, Cox et al. 2004, Magombo 2003, Wahrmund et al. 2010) and our analyses corroborate this result (Figure 1A). The Funariidae (sensu Goffinet et al. 2007), uniting the Gigaspermatales, Funariales, and Encalyptales, are resolved as a monophyletic group but without significant support (see also Goffinet & Cox 2000, Hedderston et al. 2004). The monophyly of the clade consisting of the Funariales and Encalyptales is supported by the presence of a 71kb inversion in the single-copy region of the chloroplast genome (Goffinet et al. 2007), a seemingly strong synapomorphy for the group that was recovered by our and some other phylogenetic analyses (Goffinet & Cox 2000, Cox et al. 2004).

The relationships among the three sub-classes, Funariidae, Dicranidae, and Bryidae, remain ambiguous. The Funariidae are resolved as the sister-group to the clade consisting of the Dicranidae (haplolepideae) and Bryidae (diplolepideous-alternate peristomate mosses) but lacking strong support (Figure 1A), by contrast to the inferences from multiple genes (Cox et al. 2004). Other analyses suggest that the Funariidae are most closely related to the haplolepideae: based on multiple genes (Goffinet & Cox 2000), nad5 (Beckert et al. 2001), rps4 (Goffinet et al. 2001, Hedderston et al. 2004), or nad5 plus rps4 (S.I. Figure S.2.7). However, the latter result is not confirmed by our analyses here (S.I. Figure S.2.1) which support the clade uniting the haplolepideae and Bryidae (94% p.p.). The correct resolution of these lineages remains elusive, and may well rest upon resolution of the members of the so-called 'proto-haplolepideae' (Hedderston et al. 2004), some of which were resolved as most closely related to the Bryidae in our rps4 tree (S.I. Figure S.2.1) and as most closely related to the Funariidae in previous analyses of the same gene (Goffinet et al. 2001).

The ambiguity among the three sub-classes is exacerbated when Timmia is considered. In previous analyses, Timmia has been resolved either as the sister-taxon to the Funariidae (Cox et al. 2004) or the sister-taxon to the clade uniting Dicranidae and Bryidae (Cox et al. 2004), as either the sister-taxon to (Goffinet & Cox 2000), or within the clade uniting Dicranidae and Funariidae (Goffinet et al. 2001, Beckert et al. 2001), or as the sister-taxon to the Dicranidae alone (Hedderston et al. 2004). In the rps4 analyses presented here, Timmia is strongly supported within the Encalyptaceae (S.I. Figure S.2.1), while the nuc26S analysis places it strongly in a clade with two haplolepideous taxa, Distichium ( Ditrichaceae) and Timniella ( Pottiaceae) (S.I. Figure S.2.3). Timmia’s placement within the Encalyptaceae seems most likely an artifact as it does not share the inversion in the chloroplast genome common to members of the Encalyptaceae (Goffinet et al. 2007).

The three main lineages of arthrodontous mosses, Funariidae, Dicranidae, and Bryidae are defined by distinct combinations of architectural and ontogenic features of their peristome. The architecture of Timmia’s peristome is unique but it shares critical developmental stages with the peristome of Funaria (i.e., the symmetric division in cells contributing to the inner surface of the endostome; Budke et al. 2007). With the ambiguity in the sequence of divergence of these four taxa, the fundamental question of which peristome-type
composes the ancestral architecture (Bryum-type *fide* Crosby 1980, or Funaria-type *fide* Vitt 1984) remains unanswered.

**The haplolepideae (Dicranidae)**

The monophyly of mosses with a haplolepideous peristome (Vitt 1984) has been consistently supported by phylogenetic inferences (Goffinet & Cox 2000, LaFarge *et al.* 2000, Cox *et al.* 2004, Tsubota *et al.* 2004). The circumscription of the Dicranidae as proposed by Vitt (1984) has only changed slightly, and only as the affinities of taxa with atypical, reduced, or no peristome are reassessed, suggesting that the haplolepideous architecture of the peristome (Shaw *et al.* 1989) evolved only once. The circumscription of the Dicranidae has been expanded to include taxa such as the Drummondiaceae (Goffinet *et al.* 1998), *Wardia* (Hedderson *et al.* 1999), Splachnobryaceae (Goffinet & Cox 2000, Werner *et al.* 2004) and the Ephemereaceae (Goffinet & Cox 2000, Werner *et al.* 2007), and is here further broadened to accommodate *Bryowijkia*, a genus previously included in the Hedwigiaceae (Vitt 1984) or Trachypodaceae (Vitt & Buck 1984). Should the placement of *Bryowijkia* be confirmed, then the interpretation of its peristome as being diplolepideous (Vitt & Buck 1984) must be questioned. The circumscription of the Dicranidae has also been contracted as *Bryobartramia*, long considered a member of the Pottiales (Vitt 1984), is shown to share a common ancestor with *Encalypta* (Hedderson *et al.* 2004, Goffinet *et al.* 2007).

The backbone phylogeny of the Dicranidae is poorly supported except for a) *Timmiella anomala*, *Distichium capillaceum*, and *Clastoblastella chilense* composing a clade sister to, or a grade leading to, the remainder of taxa; b) *Bryoxiphium norvegicum*, *Drummondia obtusifolia* and *Scouleria aquatica* forming the next clade or grade, and c) the monophyly of a broad assemblage including most Dicranaceae, most Ditrichaceae, Pottiaceae, Rhabdoweisiaceae, Fissidentaceae, Schistostegaceae, etc, but excluding the Grimmiaeae, Ptychomitriaceae and Leucobryaceae (Figure 1A). Whether the Dicranales as defined by Goffinet *et al.* (2009) or Vitt (1984 as Dicranineae) compose a monophyletic lineage is not clear: except for the above-mentioned taxa the Dicranales compose in most optimal reconstructions at best a paraphyletic group leading to the Pottiales but support to reject the monophyly is lacking given the overall weakness of the backbone phylogeny derived from the three loci. However, when inferences are drawn only from *rps4* and *nad5* data alone, the successive nodes of the grade are well-supported (S.I. Figure S.2.7). The polyphyly of the Dicranales is further accentuated in the latter phylogeny by the shared ancestry of the Leucobryaceae and Eustichiaeae with the Archidiales. Within the Dicranales, the circumscription of the Dicranaceae, Ditrichaceae and Rhabdoweisiaceae is in critical need of reassessment, as suggested by Stech (1999a,b) and La Farge *et al.* (2000, 2002). By contrast the circumscription of the Grimmiales as proposed by Tsubota *et al.* (2004) withstands additional sampling, although the inclusion of the Seligeriaceae is supported by inferences from cytoplasmic loci only (S.I. Figure S.2.7). The affinities of the Grimmiales remain ambiguous in all reconstructions. Similarly the sister-group to the Pottiales is still unknown.

Overall, the lack of resolution of the successive cladogenic events marking the macroevolutionary history of the Dicranidae suggests that diversification was rapid following the evolution of the haplolepideous peristome. Indeed, an overall slow substitution rate in mosses provides an inadequate explanation for the short branches defining putative lineages considering that more recent cladogenic events are seemingly defined by rather unambiguous synapomorphies (e.g., the well supported Pottiales, Leucobryaceae, or Grimmiaeae). If a rapid radiation occurred, resolving the branching order within the Dicranidae may required intensive sampling (i.e., 10 or more loci), as was needed for the resolution of the relationships within the Saxifragales for example (Jian *et al.* 2008). Resolution of the relationships and the circumscription of major lineages within the haplolepideous mosses is essential to critically assess the phylogenetic and ecological significance of transformations of gametophytic and sporophytic characters in particular those linked to transitions to extreme habitats.
The acrocarpous diplolepideous-alternate peristome mosses

Within the unambiguously monophyletic Bryidae (i.e., mosses with a diplolepideous-alternate peristome mosses), the superorder Bryanae (sensu Goffinet et al. 2009 — Bryanae and Rhizogonianae but excluding Hypnodendrales sensu Buck & Goffinet 2000) composed of essentially acrocarpous taxa is a grade from which the pleurocarpous mosses (Hypnanae) are derived. Although the paraphyly of the Bryidae has long been suspected (e.g. Cox & Hedderson 1999, Newton et al. 2000), the sister-group to the pleurocarpous mosses remains unknown. The combined analyses presented here place the Orthotrichales as the well-supported sister-group to the ‘rhizogoniaceous grade’ of (with successive branching) Orthodontiales, Rhizogoniales, Aulacomniales, Hypnodendrales, and the pleurocarpous mosses (Bell et al. 2007). This order of branching among the groups is mainly due to support from the mitochondrial nad5 gene (S.I. Figure S.2.2) - the chloroplast rps4 data (S.I. Figure S.2.1) provide little supporting or contradictory evidence, as noted by Bell et al. (2007). Our analyses of the nuc26S gene data indicate highly unusual groupings, for instance, the Ptychomniales and Aulacomniales being more closely related to the Bryales (S.I. Figure S.2.3). These latter relationships most likely reflect analytical artifacts, perhaps a failure to adequately model compensatory changes in stem regions of the 26S rRNA molecule (Dixon & Hillis 1993). Nevertheless, these results need to be confirmed with additional data from the nuclear genome.

Bartramiales and Hedwigiales are resolved as the first diverging lineage within the Bryanae (Figure 1A). However, the Splachnales were excluded from the tree due to incongruence among partitions — rps4 placed the Splachnales in this basal-most position, but the nuc26S resolved them nested within the Hypnales (again most likely an analytical artifact). Support for the basal-most position of Splachnales within the Bryanae is weak: 83% p.p. (rps4: S.I. Figure S.2.1; 71–77 p.p. Bell et al. 2007) and contradicted by the nad5 data (80% p.p; see also Cox & Hedderson 1999, Cox et al. 2004). Similarly, support for the sister-relationship between Bartramiales and Hedwigiales is lacking, though not statistically contradicted, in our analyses (see also Cox et al. 2000, Goffinet et al. 2001, Bell et al. 2007). Yet, remarkably, nuc26S provides very strong support for the grouping of all other Bryanae to the exclusion of the Bartramiales and Hedwigiales (S.I. Figure S.2.3), possibly artifactually, but congruent with the multigene analysis of Cox et al. (2004).

The pleurocarpous mosses

Through evolutionary time, the transfer of the female sex organ to the apex of short specialized branches marks the origin of the Hypnanae, the most diverse lineage of mosses, comprising approximately 6,000 species. For much of the 20th century these species were accommodated into three main lineages, recognized at different ranks, but equivalent to the Hookeriales, Hypnales, and Leucodontales (Vitt 1982, 1984). The pioneering study by Hedenäs (1994) noted that these concepts were unlikely to withstand critical cladistic analysis of morphological characters, and the first inferences from DNA sequence data (DeLuna et al. 1999) lent credence to that hypothesis. The polyphyly of the Leucodontales was demonstrated by DeLuna et al. (1999) and Buck et al. (2000) and the order has not been recovered as a monophyletic entity since (Tsubota et al. 2004, Ignatov et al. 2007). Its families have been transferred to the Hypnales, except for the Ptychonomyiaceae, which are now considered the sister lineage to all other pleurocarpous mosses (Buck et al. 2004). The circumscription of the Hookeriales, the sister-group to the Hypnales, has remained unchanged since Vitt (1984) although the number of recognized families has increased from three to seven (Buck et al. 2004). In the analyses presented here the Hypopterygiaceae are resolved within the Hypnales rather than the Hookeriales, as previously suggested by Pedersen & Newton (2007), but none of the nodes separating the Hypopterygiaceae from the Hookeriales are well-supported, and hence a shared ancestry with other Hookeriales as proposed by Buck et al. (2004) cannot be rejected. In contrast to the well-supported monophyly of the Hookeriales s. str., a unique shared ancestry for all Hypnalean taxa remains doubtful (Figure 1B; Tsubota et al. 2004, Olsson et al. 2009a). The lack of synapomorphies for the Hypnalean taxa may be a consequence of their rapid diversification after divergence from the Hookeriales (Shaw et al. 2003b) that may have followed the diversification of angiosperms and the development of the angiosperm forest ecosystems (Pedersen & Newton 2007). A rapid diversification coupled with a low rate of mutations (Stenøien 2008) may
also account for the conspicuous lack of robustness throughout the backbone of the optimal phylogeny of pleurocarpous mosses inferred from all three loci, a weakness highlighted earlier by Goffinet et al. (2001) and Tsubota et al. (2004) and observed again most recently by Olsson et al. (2009b). Regardless of the character source and the taxon sampling, few well-supported lineages of hypnalean taxa are recovered and these lineages rarely match the families defined by Vitt (1984) or contemporary workers. The phylogenetic inference presented here includes several large robust monophyletic lineages in the Hypnales, that can be roughly compared to the Pterobryaceae, the Meteoriaceae, the Brachytheciaceae (Figure 1B), Sematophyllaceae s. lat. (including the Pylaisiadelphaceae), and the Amblystegiaceae (Figure 1C). Except maybe for the extensively studied Meteoriaceae (Quandt et al. 2004, Huttunen & Quandt 2007) and Lembophyllaceae (Quandt et al. 2009), these families in the sense of Goffinet et al. (2009) are polyphyletic, due to the scattered distribution of some discrete outliers. One other clade includes most Neckeraeae, but not the type genus, a hypothesis contrary to the recent study by Olsson et al. (2009) wherein inferences were drawn from more variable loci (but see also Olsson et al. 2009b). Noteworthy here may be the resolution of the monogeneric Echinodiaceae, a taxon not sampled by Olsson et al. (2009) within the Neckeraeae (Figure 1C). Unlike most families of pleurocarpous mosses, the Thuidiaceae share a common ancestor, although it is unique only if the Helodiaceae and Haplocladium (Leskeaceae) are accepted within a broader concept of the family. A core group of Cryphaeaceae composes a robust lineage to the exclusion of Dendroalsia and Pilotrichopsis (Figure 1B).

Perhaps the most notable result of the our phylogeny is the widely polyphyletic nature of the Hypnaceae, a family traditionally defined by the ecostate leaves, elongate-linear cells, and a well developed double peristome with a high endostomial membrane. Vitt (1982) had doubted that the circumscription of the family would withstand critical examination: a conjecture that has been borne out by the phylogenetic analysis presented here (Figure 1B&C) and earlier (Tsubota et al. 2004).

Phylogenetic diversity

Biodiversity can have multiple meanings in different contexts and can consequently be estimated using different metrics depending on the purpose. Phylogenetic diversity (PD) quantifies the amount of evolutionary change (branch lengths on a phylogenetic tree) that separates taxa, and can provide information that complements diversity estimates based on numbers of species or some other taxonomic unit (Faith 1992, 1994). We observed that PD and taxic diversity give highly correlated estimates of how biodiversity is partitioned among some of the major clades of mosses; indeed correlation coefficients were greater than 0.9 between estimates based on loci from the three different genomes (plastid, mitochondrial, nuclear), and between these molecular estimates and numbers of genera. One implication of this consistency is that the generic level in the Linnean taxonomic hierarchy for mosses is not as arbitrary as one might perhaps suggest. That is, a genus of Funariales, for example, encompasses about the same amount of phylogenetic diversity as does a genus of Bartramiaceae.

PD estimates can in some cases provide insight into the tempo and mode of evolutionary diversification. In the peatmosses (Sphagnum), the Neotropics contain an extremely high level of species richness, but much less PD than predicted from those species numbers (Shaw et al. 2003a). One possible explanation from that observation is that species richness of Neotropical peatmosses reflects a recent radiation. Similarly, the lower than predicted (from taxic diversity) levels of PD within hypnalian pleurocarpous genera may be related to a relatively recent and sudden radiation of this clade (Shaw et al. 2003b). Higher than predicted levels of PD in the Dicranidae could, conversely, reflect extinction, leading to more distinct; i.e., relictual, genera. Other explanations such as differences in substitution rates among clades cannot be ruled out, but it is noteworthy that the patterns are highly correlated across the three genomes, supporting a more general interpretation. Phylogenetic estimates of biodiversity do not substitute for more traditional taxon-based metrics, but can generate hypotheses about evolutionary processes.
Concluding remarks

Our analyses highlight the difficulty of obtaining robustly supported groupings of the main lineages of mosses. Despite the antiquity of the main lineages (Newton et al. 2007), the overall rate of moss molecular evolution appears to be rather slow (Stenøien 2008) suggesting that many loci will most likely be required to confidently establish some major divergences. At a shallower taxonomic level, the many instances of polyphyly among the families of haplolepidous and pleurocarpous mosses re-inforce the need to address phylogenetic hypotheses within a broader context than those defined by traditional families.

Obtaining statistically robust resolutions of the relationships, firstly, among the nematodontous taxa, Oedipodium, Buxbaumia, and the arthrodontous mosses, and, secondly, among the Funariidae, Dicranidae, and Bryidae, are the two most difficult problems currently facing moss phylogenetics, with respect to deep evolutionary divergences. The conflicting resolutions of the Funariidae plus Dicranidae versus the Dicranidae plus Bryidae hypotheses may be indicative of an evolutionary event resulting in authentic phylogenetic conflict (such as an early hybridization event between two of these lineages); the cytoplasmic phylogenetic signal for the former relationship is strong, but always overruled (although without support) when the nuclear rRNA loci are included. Yet, the alternative explanation of this conflict being an artifact resulting from model mis-specification during the phylogenetic analyses needs to be disproved before the implications of a hybridization event deep in moss evolution can be fully assimilated.

Although some crucial nodes in the moss phylogenetic tree remain poorly resolved and therefore of minimal value for the reconstruction of ancestral character states, the completion of the first entire moss genome (Physcomitrella: Rensing et al. 2008), and the ability to characterize transcriptomes at particular ontogenetic stages, provides an opportunity for a complementary experimental approach to moss systematics. The resolution of gene networks underlying the development of specific traits should permit an assessment of the homology among distinct peristome types and among the transitions of gametangial production from the stem to the branches that occurred during the evolution of pleurcarpy. Solving these puzzles will be essential to understanding the innovations that mark the major cladogenic events in the evolution of the mosses.

Acknowledgements

We wish to thank William Buck, Ingo Holz, Blanka Shaw, and Neva Hax for their contributions to the project. We also gratefully acknowledge the curators of the botanical gardens of Missouri (MO) and New York (NY) for allowing us access to their collections. This project was supported by National Science Foundation grants DEB-0089131 to AJS and DEB-0089633 to BG.

References

Bridel-Brideri, S.E. (1826–1827) Bryologica Universa seu Systematica ad Novam Methodum Dispositio, Historia et


